Attachment C

I. Methodology Used For Finding Duplicate Records In The T&N Claims Database

To identify potential duplicate records in the T&N Claims Database, NCI staff looked at the consolidated data containing claimant and claim information, such as the settlement amount (net expense), disease type, status, and the filing and other dates for the claim. For the analysis of duplicate records, NCI staff looked at entries for claimants with the same first and last names, and social security number; 28,093 records met these criteria.

A. Claimants with more than two claims

Within this group were three claimants, each with four claims listed with the same disease. These three claimants each had three claims with a "Settled/Paid" status and one claim with an "Open" status. NCI staff retained the "Settled/Paid" claim with the largest settlement amount and identified the other three claims as duplicates (a total of nine duplicates).

The remaining 28,081 claims included 21 claimants, each with three claims filed (a total of 63 claims). Of these claims, NCI staff identified 23 claims with another claim filed for the same disease or a disease less severe than the already settled claim. These claims were identified as duplicates. NCI staff applied the following decision rules to de-duplicate these claims.

For claims with same disease:

1. Eliminate those with the status "Closed" where there is a "Settled/Paid" claim for the same disease.
2. Keep earliest claim, if more than one has the same status.
3. Eliminate claims in the following order of status:
    a.   "Settled But Not Documented," if another claim is "Closed" or "Settled/Paid."
    b. "Open," if another claim is "Dismissed With Prejudice" or "Settled/Paid."
    c. "Closed," if another claim is "Settled/Paid."
4. Eliminate "Open" claims with Unknown diseases for claimants with more than two remaining claims.

B. Claimants with two claims

The number of claims in this group was 28,018. Within this group, 20,852 had the same disease and 7,166 had either a different or a missing disease (we used the defense disease category to match the disease types).

NCI staff applied the following rules to de-duplicate the file, identifying 9,337 duplicate claims.

For claims with same disease

1. Eliminate later claim when both claims have the same status (for "Settled/Paid" status we looked at expense date and for all other dates we used the earliest of filing, served, and received dates).

2. Eliminate if there is a "Settled But Not Documented" claim earlier than the "Settled/Paid" claim, where the "Settled/Paid" claim had a positive payment amount.

3. Eliminate "Settled/Paid" claim if there is a zero payment on a "Settled But Not Documented" claim.

4. If claims have different statuses:

   a. Eliminate earlier claim if:

      i. "Closed" and later claim is "Dismissed With Prejudice," "Settled/Paid," or "Open."

      ii. "Open" and later claim is "Dismissed With Prejudice" or "Settled/Paid."

      iii. "Dismissed With Prejudice" and later claim is "Settled/Paid" (with a positive payment amount)

      iv. "Settled/Paid" (with a zero payment) and later claim is "Dismissed With Prejudice."

   b. Eliminate later claim if:

      i. "Open," and earlier claim is "Dismissed With Prejudice" or "Settled/Paid."

      ii. "Settled/Paid" (with a zero payment), and earlier claim is "Dismissed With Prejudice."

      iii. "Closed" and earlier claim is "Settled/Paid" (with a positive payment amount).

For the 7,166 claims where the defense disease was different or missing, NCI staff applied the following decision rules to identify the duplicate claims.

For claims with different diseases

1. Eliminate claim with less severe disease if both are "Open."

2. Eliminate claim with less severe disease if "Open" and other claim is "Settled/Paid."

3. Keep both claims under all other circumstances.

For claims with unknown diseases

1. Eliminate Unknown disease claim if both claims have the same status.

2. Eliminate "Open" disease claim if a "Settled/Paid" claim (with a positive payment amount) is for a malignant disease.

3. Eliminate "Open" nonmalignant claim if Unknown disease claim is "Settled/Paid" (with a positive payment amount).

4. Keep both claims under all other circumstances.

Applying these decision rules, NCI staff found 3,467 duplicates within those that had a different or a missing disease.

For claimants with more than one filed claim, NCI staff found 12,836 duplicate claims using the decision rules listed above. The duplicate claims were identified with a duplicate flag[1] in the data base and merged back into the T&N consolidated data.

II. Procedures For Matching T&N Claimants To Manville Trust Claimants

NCI staff attempted to fill in relevant information about T&N asbestos claimants that was missing from the T&N Claims Database by matching each T&N claimant to the identical claimant in the Manville Trust asbestos claims database ("Manville database") and extracting the missing information from the latter.[2] NCI staff used the Manville database dated September 2004, the most recent available. Because the claimant name fields were to be used as part of our matching criteria and because there might be slight typographical differences in the recording of these names between the T&N and Manville databases, NCI staff first made minor edits in the name fields so that names would be recorded consistently (e.g., deleting blank spaces within a name).

After these edits, NCI staff developed sets of criteria to use to match the T&N and Manville claimants. Ultimately NCI staff used 12 sets of criteria that successfully matched some T&N claimants. NCI staff implemented these sets of criteria in descending order of their stringency. Thus, NCI staff started with the set of criteria likely to produce the most reliable match between a T&N and Manville claimant. Only those T&N claimants that were not matched to a Manville claimant by this set of criteria were subjected to a match based on the next most stringent set of criteria. This sequence continued until all 12 sets of matching criteria were exhausted. The 12 sets of matching criteria, in the order they were implemented, are as follows, as well as the number of T&N claimants matched by the set:

#1: First names, last names, social security numbers, and birth dates match, and the death dates either match or at least one of them is blank[3]
203,111

---

[1] NCI staff added a new field identifying the duplicates as "dup_flag=1."

[2] NCI staff only extracted the Manville disease information for Manville claims that were not still being processed, because the listed disease for claims that were not final might not reflect the disease ultimately decided upon.

[3] Identical claimants in the T&N and Manville databases need not have the same death date because the claimant could have been alive when his/her claim was resolved with one of the defendants, but died before his/her claim was resolved with the other defendant. Thus, as long as the death dates were not both present and inconsistent, a blank death date in one did not indicate there was no match.

#2: First names, last names, social security numbers, and birth years match, and the death dates either match or at least one of them is blank
7,067

#3: First names, last names, social security numbers, and birth dates match, and the death years either match or at least one of them is blank
1,121

#4: Last names, social security numbers, birth dates, and first five characters of first names match, and the death dates either match or at least one of them is blank
1,475

#5: First names, social security numbers, birth dates, and first five characters of last names match, and the death dates either match or at least one of them is blank
1,511

#6: First names, last names, social security numbers match, and the birth and death dates either match or at least one of them is blank
50,335

#7: First names, last names, birth dates, and first six characters of social security numbers match, and the death dates either match or at least one of them is blank
2,136

#8: First names, last names, birth dates, and last six characters of social security numbers match, and the death dates either match or at least one of them is blank
1,970

#9: First names, last names, birth dates, and first five characters of social security numbers match, and the death dates either match or at least one of them is blank
782

#10: First names, last names, birth dates, and first and last three characters of social security numbers match, and the death dates either match or at least one of them is blank
783

#11: Social security numbers match
16,743

#12: First names, last names, and first six characters of social security numbers match, and the birth and death dates either match or at least one of them is blank
649

C-4

Of the T&N claimants, a total of 75.0% were matched to a Manville claimant, and 89.4% of the T&N claimants with social security numbers were matched to a Manville claimant.[4]

---

[4] Because 16.5% of T&N claimants were missing social security numbers and essentially all of them also were missing birth dates, these claimants lacked sufficient information to reliably match to Manville claimants.

C-5